

Корреляционный анализ и равенство средних

В эксперименте измерено 6 значений (var0, var1, var2, var3, var4, var5) для 30-40 объектов. Данные записаны в файл `data1-vX.csv`. (число после буквы `v` в имени файла – номер варианта).

Предполагаем, что var5 зависима от остальных (var0-var4).

Для набора данных 1 (`data1-vx.csv`):

1. Построить матрицу корреляции.
2. Проверить гипотезы о значимости каждого коэффициента каждой величины с пятой (var5).
3. Построить диаграмму рассеивания для var5 и двух других величин: одну для самой слабой корреляции, другую для наиболее сильной.
4. Записать уравнение линейной регрессии для var5 и другой величины, корреляции с которой наибольшая.
5. Оценить предсказательную способность уравнения регрессии с помощью коэффициента детерминации R^2 .

В эксперименте измерялись значения двух величин для нескольких десятков объектов. Были получены зависимые выборки. Для этого набора данных 2 (`data2-vX.csv`):

1. Построить диаграммы размаха для выборок.
2. Найти 95% доверительный интервал для средних значений выборок, полагая что генеральные совокупности имеют нормальное распределение.
3. Проверить гипотезу о равенстве средних значений для двух выборок.
4. Проверить гипотезу о равенстве двух выборок используя непараметрический критерий

Вопросы

1. Что такое выборка? Что такое генеральная совокупность?
2. Что такое корреляция?
3. Что такое коэффициент корреляции (линейный)?
4. Что такое статистическая гипотеза?
5. Что такое p-value? Как это значение используется для проверки гипотез?
6. Почему нужно проверять гипотезу о равенстве коэффициента корреляции генеральной совокупности нулю?
7. Что такое диаграмма размаха?
8. Что такое матрица корреляции?

Ссылки

1. Гмурман, В. Е. Теория вероятностей и математическая статистика : учебник для прикладного бакалавриата / В. Е. Гмурман. — 12-е изд. — Москва : Издательство Юрайт, 2019. — 479 с. — (Бакалавр. Прикладной курс). — ISBN 978-5-534-00211-9. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://biblio-online.ru/bcode/431095> (дата обращения: 21.12.2019)
2. Яковлев, В. Б. Статистика. Расчеты в Microsoft Excel : учебное пособие для вузов / В. Б. Яковлев. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2021. — 353 с. — (Высшее образование). — ISBN 978-5-534-01672-7. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/471895> (дата обращения: 18.12.2021).
3. https://colab.research.google.com/drive/1JUW43GBoo1DNczcpCohe7sjee_EHuxbo – пример на Python